FLA-Net: 面向反无人机场景的频域增强与低时延检测框架

刘佳祺[1]，张顺香[1,2,3†]，张正佳[1]，王 杰[4]

(1. 安徽理工大学 计算机科学与工程学院，安徽 淮南 232001；
2. 淮南师范学院 计算机学院，安徽 淮南 232038；
3. 合肥综合性国家科学中心人工智能研究院，合肥 230000；
4. 安徽理工大学 安全科学与工程学院，安徽 淮南 232001)

摘 要: 基于深度学习的固定视角可见光反无人机检测能够为低空安全预警提供目标外观、位置和语义信息. 然而, 当远距离小尺度无人机靠近视场边界时, 可观测余量有限、结构线索较弱, 导致检测质量与响应时效的平衡问题更加突出. 为此, 提出频域-时延协同优化检测框架 FLA-Net, 并设计视场边界风险比, 用于评估边缘目标在单帧推理期间接近或超出视场边界的潜在风险. 该框架围绕 YOLO26n 设计自适应频域重分配模块和层级低时延特征聚合模块, 前者通过频域分解和子带重标定增强边缘小尺度目标的高频细节表达, 后者通过轻量通道分组递推聚合降低特征聚合开销, 并利用多尺度上下文提升特征表达能力. 在 DUT Anti-UAV 数据集上的固定视角可见光检测实验表明, FLA-Net 在较低计算开销下兼顾检测质量与低时延推理, 并具有较低的视场边界风险.

关键词: 深度学习；固定视角反无人机检测；小目标检测；视场边界风险；频域特征增强；低时延推理

中图分类号: TP391.41 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2026.0189

引用格式: 刘佳祺, 张顺香, 张正佳, 等. FLA-Net: 面向反无人机场景的频域增强与低时延检测框架 [J]. 控制与决策.

FLA-Net: A frequency-domain enhancement and low-latency detection framework for anti-UAV applications

LIU Jia-qi[1]，*ZHANG Shun-xiang*[1,2,3†]，*ZHANG Zheng-jia*[1]，*WANG Jie*[4]

(1. School of Computer Science and Engineering，Anhui University of Science and Technology，Huainan 232001，China；2. School of Computer Science，Huainan Normal University，Huainan 232038，China；3. Artificial Intelligence Research Institute of Hefei Comprehensive National Science Center，Hefei 230000，China；4. School of Safety Science and Engineering，Anhui University of Science and Technology，Huainan 232001，China)

Abstract: Deep learning-based fixed-view visible-light anti-UAV detection can provide target appearance, location, and semantic information for low-altitude security early warning. However, when distant small-scale UAVs approach the field-of-view boundary, the limited observation margin and weak structural cues make the balance between detection quality and response timeliness more challenging. To address this issue, this paper proposes FLA-Net, a frequency-domain enhancement and low-latency detection framework, and introduces a field-of-view boundary risk ratio to evaluate the potential risk of boundary targets approaching or moving beyond the field-of-view boundary during single-frame inference. Built upon YOLO26n, FLA-Net incorporates an adaptive frequency redistribution module (AFRM) and a hierarchical low-latency feature aggregation module (HLFA). AFRM enhances the high-frequency detail representation of small-scale targets near the field-of-view boundary through frequency decomposition and subband recalibration. HLFA reduces feature aggregation overhead through lightweight channel-grouped recursive aggregation and improves feature representation through multi-scale context modeling. Experiments on the DUT Anti-UAV dataset under the fixed-view visible-light detection setting demonstrate that FLA-Net achieves a favorable balance between detection quality and low-latency inference with low computational complexity and exhibits a low field-of-view boundary risk.

Keywords: deep learning； fixed-view anti-UAV detection； small object detection； field-of-view boundary risk； frequency-domain feature enhancement；low-latency inference

收稿日期: 2026-03-10；**录用日期:** 2026-08-10.
基金项目: 国家自然科学基金面上项目 (62476005, 62076006).
责任编委: 王琦.
†通信作者. E-mail: sxzhang@aust.edu.cn.

0 引 言

随着无人机技术的快速发展和普及, 低空空域安全与防御需求日益突出[1]. 在要地防空、机场净空区及关键基础设施防护等低空安全场景中, 反无人机系统需要稳定的目标感知能力[2]. 雷达等专用探测手段为反无人机系统提供了重要的感知基础[3]. 同时, 在视距飞行、空域监管和目标可视确认等实际约束下[4], 可见光检测能够提供目标外观、位置和语义信息, 因而成为反无人机探测链路中的重要视觉前端[5]. 近年来, 面向低空安全场景的视觉感知研究持续受到关注[6]. 对于固定视角可见光反无人机检测而言, 检测结果质量与响应速度会影响后续跟踪、识别与决策环节的输入可靠性.

在后续跟踪与主动观测环节中, 目标可见性和实时响应是维持稳定观测的重要因素[7]. 当目标接近视场边界或观测更新不及时, 状态估计与后续决策容易受到目标丢失或输入失配影响. 对于固定视角可见光检测而言, 检测结果是后续处理获得目标观测的起点, 因此有必要将对视场内目标观测有效性与检测响应时效性的关注提前至检测阶段. 连续跟踪和主动观测侧重目标轨迹状态的维持, 固定视角反无人机检测则侧重在逐帧处理中及时、可靠地检出边缘目标, 为后续环节提供有效观测输入.

近年来, 深度学习显著推动了视觉反无人机检测技术发展[8]. 相关方法通常围绕复杂背景中的小目标特征增强[9]、轻量化网络设计[10]和多尺度特征融合[11]展开, 以提升无人机目标检测性能. 面向无人机航拍图像的小目标检测, 近期研究从多尺度特征表达、特征增强聚合和轻量化检测头等方面开展改进, 进一步体现了检测精度与计算效率协同优化的重要性[12]. 反低空无人机目标检测研究表明, 复杂背景下的小目标漏检问题及轻量化约束下的检测性能仍是视觉检测前端需要关注的重点[13]. 在通用视觉特征学习中, 已有研究将二维 Haar 小波分解与注意力机制结合, 利用多频带信息增强卷积神经网络的特征表征能力[14]. 与此同时, 视觉反无人机研究已从目标检出逐步延伸到联合跟踪、目标描述和语义理解等任务链路, 检测结果在后续任务中的输入作用受到更多关注[15]. 现有检测性能评价多以整体检测精度、平均推理时延或帧率作为主要指标, 能够反映模型的全局性能, 但难以进一步刻画视场边界条件下检测结果对后续观测输入的支撑能力. 因此, 在整体精度、平均推理时延或帧率等全局指标之外, 补充面向视场边界条件的评价维度, 有助于从场景约束出发评价检测前端的任务适应性.

为回应上述评价需求, 本文从评价建模和检测方法设计两个方面展开研究: 一方面, 构建面向固定视角反无人机检测的边界风险评估模型, 用于分析边缘目标在视场边界余量与推理时延共同作用下的观测风险; 另一方面, 以 YOLO26n 为基线进行方法设计, 以提升边缘小尺度目标检测质量, 并兼顾模型轻量化与实时响应. 基于上述思路, 本文主要贡献如下:

1) 设计视场边界风险比 (field-of-view boundary risk ratio, FBRR), 结合目标距视场边界的像素余量、目标尺度与推理时延, 评估边缘目标在单帧推理期间接近或超出视场边界的潜在风险; 进一步结合检测状态和预测置信度构建单目标级视场边界联合风险指数 R_i, 为固定视角反无人机检测任务提供补充评价依据.

2) 提出频域-时延协同优化检测框架 FLA-Net: 在骨干特征提取部分设计自适应频域重分配模块 (Adaptive Frequency Redistribution Module, AFRM), 通过频域分解、子带重标定和受限残差融合增强边缘小尺度目标的高频细节表达; 在颈部特征聚合部分设计层级低时延特征聚合模块 (Hierarchical Low-latency Feature Aggregation Module, HLFA), 通过轻量通道分组递推聚合降低特征聚合开销, 并利用局部-区域-全局上下文建模提升多尺度特征表达.

3) 在 DUT Anti-UAV 可见光检测数据集上开展整体检测性能、边缘目标图像组检测性能和 ONNX Runtime 后端推理时延实验, 并进行基于 FBRR、检测状态和预测置信度的视场边界风险联合分析, 验证 FLA-Net 能够在较低计算开销下兼顾检测质量与低时延推理, 并具有较低的视场边界风险.

1 任务描述与问题设定

本文研究固定视角可见光反无人机检测任务中的边缘目标观测问题. 该任务以固定视场图像为输入, 由检测器输出无人机目标框及其预测置信度, 为后续跟踪、识别与告警等处理环节提供观测输入. 对于靠近视场边界的目标, 其检测框中心到视场边界的像素余量有限, 检测结果的输出时效可能影响后续观测输入的可靠性.

边缘目标的观测条件主要受视场边界余量、目标图像尺度和检测推理时延共同影响. 视场边界余量反映检测框中心当前位置与视场边界之间的距离; 目标图像尺度用于表征目标在图像平面中的可观测尺寸, 并作为后续风险计算中的尺度因素; 检测推理

时延则表征模型完成一次前向执行所需的时间. 当边界余量较小或推理时延较大时, 边缘目标在单帧推理期间接近或超出视场边界的潜在风险会增加. 因此, 有必要将上述因素纳入统一评估框架, 建立面向固定视角检测任务的边界风险评估模型.

为建立该模型, 本文作如下约定与假设:

(1) 约定检测器输入为固定视角可见光图像, 输出为无人机目标框及其预测置信度, 目标在图像中的位置由检测框中心表示;

(2) 约定在给定图像尺寸下, 目标中心到最近图像边界的距离表示视场边界余量, 目标图像尺度由目标框宽、高中的较大值表示;

(3) 考虑到检测标注未直接提供目标物理速度和物理尺寸, 本文依据主流民用无人机公开规格选取代表性物理速度 v 和物理基准尺寸 L;

(4) 为估计目标在单帧推理期间相对于最近视场边界的位置变化, 采用局部匀速并朝向最近视场边界运动的保守假设;

(5) 约定检测器完成一次前向推理所需时间由固定硬件和固定推理后端实测获得, 并与目标图像尺度、代表性物理速度、物理基准尺寸和视场边界余量共同用于视场边界风险比 FBRR 的计算.

2 方法介绍

2.1 边界风险评估模型

基于上述任务描述与问题设定, 本文建立边界风险评估模型, 用于分析固定视角检测任务中边缘目标在单帧推理期间接近或超出视场边界的潜在风险. 该模型以检测框中心到视场边界的像素余量表征当前边界余量, 并结合目标图像尺度、代表性物理速度和检测器模型执行时延估计单帧推理期间的图像平面位置变化, 进而定义视场边界风险比 FBRR. FBRR 不用于精确预测目标三维运动轨迹, 而是在缺少相机标定、目标深度和真实运动速度标注条件下, 对不同检测器在相同实验设置下的边界观测风险进行相对比较. 后续比较以同一组边缘真实目标为计算对象, 检测器差异通过各自的模型执行时延 τ 引入.

首先, 根据目标在图像中的位置定义其视场边界余量. 设输入图像宽、高分别为 W 和 H, 真实目标框中心坐标为 (c_x, c_y). 目标中心到图像四条边界的最小距离定义为视场边界余量 d_e, 即

$$d_e = \min(c_x, W - c_x, c_y, H - c_y). \tag{1}$$

该量越小, 表示检测框中心越接近视场边界, 其到最近边界的像素余量越有限.

其次, 根据目标尺度、代表性物理速度和检测器模型执行时延估计单帧推理期间的位置变化. 设真实目标框宽、高分别为 w 和 h, 目标图像尺度为 $l = \max(w, h)$; 无人机物理基准尺寸为 L, 代表性物理速度为 v, 检测器单帧模型执行时延为 τ(计算时以 s 为单位). 在局部匀速假设下, 本文采用目标图像尺度与物理基准尺寸之比近似表示图像平面中的尺度关系, 并以朝向最近视场边界的运动作为保守近似, 得到目标中心在单帧推理期间的估计位置变化量:

$$\Delta p = v\tau \frac{l}{L}. \tag{2}$$

其中, Δp 表示估计的像素位置变化量. 该估计主要反映物理速度、推理时延和目标图像尺度共同作用下的相对位移规模.

最后, 将单帧推理期间的估计位置变化量与视场边界余量进行比较, 定义视场边界风险比 (field-of-view boundary risk ratio, FBRR) 为

$$\rho = \frac{\Delta p}{d_e} = \frac{v\tau l}{L d_e}. \tag{3}$$

式中要求 $d_e > 0$; 当 $d_e \leqslant 0$ 时, 令 $\rho = +\infty$, 表示边界风险达到极限. 由式可见, 在代表性物理参数一致的条件下, ρ 随推理时延和目标图像尺度增大而增大, 随视场边界余量增大而减小. 因此, ρ 越大, 表示目标在一次模型前向执行期间消耗当前边界余量的比例越高, 边缘目标接近或超出视场边界的潜在风险越大. 当 $\rho \geqslant 1$ 时, 估计位置变化量达到或超过当前视场边界余量, 表示目标中心在该保守假设下可能到达或越过最近视场边界.

2.2 FLA-Net 总体架构

面向固定视角反无人机检测任务的轻量化与低时延需求, 检测网络需要权衡特征分辨率、定位质量与计算复杂度. 所提出 FLA-Net 以 YOLO26n 为基线模型, 保留基线模型的 $P_3 \sim P_5$ 多尺度检测输出, 不额外设置计算开销较高的 P_2 高分辨率检测分支. 框架整体由骨干网络、颈部网络和检测头组成, 其特征流转过程如图 1 所示.

在固定视角可见光反无人机检测中, 远距离及视场边缘小尺度无人机的细节结构容易在连续下采样过程中被削弱; 同时, 颈部网络中的多级上采样、拼接和卷积操作会增加前向传播开销. 为此, FLA-Net 从频域细节保持和低时延特征聚合两个方面对 YOLO26n 进行改进.

在骨干网络阶段, 本文在浅中层特征提取位置设计自适应频域重分配模块 (Adaptive Frequency

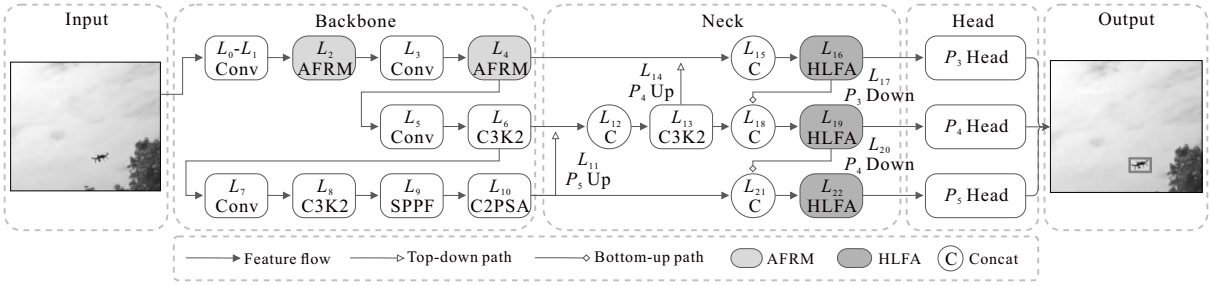

图 1 FLA-Net 总体架构

Redistribution Module, AFRM), 通过频域分解、子带重标定和受限残差融合增强小尺度目标相关的边缘、轮廓和纹理细节表达. 在颈部网络阶段, 于 P_3、P_4 和 P_5 特征聚合输出位置布置层级低时延特征聚合模块 (Hierarchical Low-latency Feature Aggregation Module, HLFA). 其金字塔聚合路径通过轻量通道分组递推聚合降低特征聚合开销, 检测侧细化路径利用局部、区域和全局上下文建模提升多尺度预测特征的表达能力.

最终, FLA-Net 仍在 P_3、P_4 和 P_5 三个尺度上输出分类置信度与边界框回归结果, 保持基线检测接口和后处理流程不变. 通过上述设计, FLA-Net 在不增加检测尺度的条件下, 提升边缘小尺度目标细节表达和多尺度定位能力, 并兼顾低时延推理需求.

2.3 自适应频域重分配模块 (AFRM)

在固定视角可见光反无人机检测中, 远距离及视场边缘小尺度无人机通常仅占据较少像素, 其可分辨性主要依赖边缘、轮廓和纹理等高频细节. 随着网络连续下采样, 小尺度目标的高频响应容易逐层减弱; 若直接放大高频分量, 又可能同步增强背景纹理和局部噪声. 为此, AFRM 以空间主路径和频域细化路径为主体, 并设置全局调节分支, 在保持基础空间表征的同时对不同频率分量进行自适应调整. AFRM 的整体结构如图 2 所示.

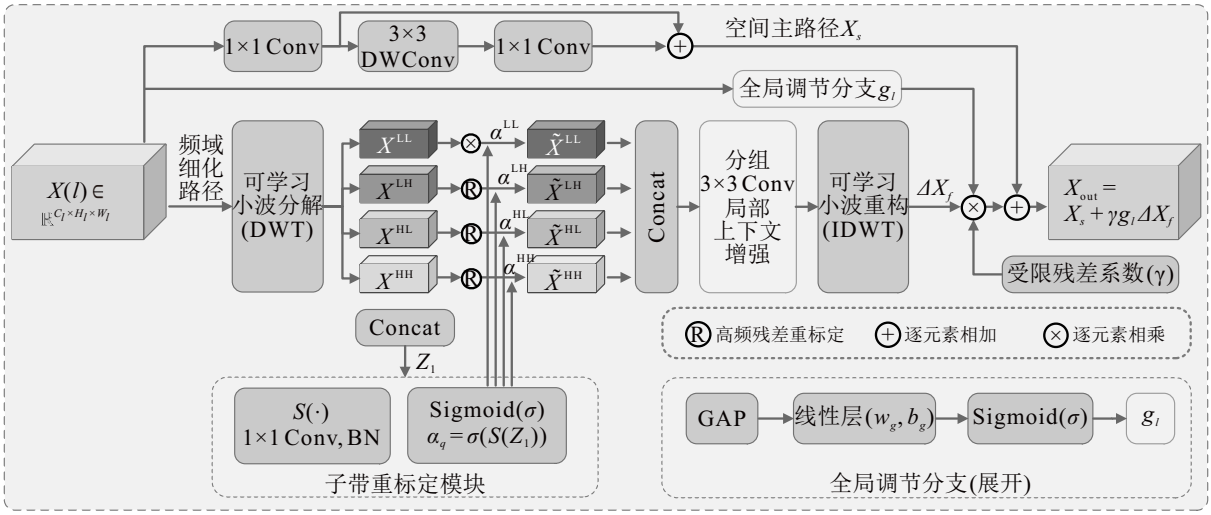

图 2 自适应频域重分配模块 (AFRM)

设第 l 层输入特征为 $X^{(l)} \in \mathbb{R}^{C_l \times H_l \times W_l}$. 空间主路径首先通过 1×1 卷积完成通道映射, 随后利用由 3×3 深度卷积、1×1 逐点卷积和残差连接构成的轻量残差单元提取局部空间信息, 得到空间表征 X_s. 全局调节分支利用全局平均池化压缩输入特征的空间信息, 得到通道统计向量 $\boldsymbol{z}_l = \mathrm{GAP}(X^{(l)})$, 并经线性映射和 Sigmoid 函数生成全局调节系数:

$$g_l = \sigma(\boldsymbol{z}_l^{\mathrm{T}} \boldsymbol{w}_g + b_g). \tag{4}$$

其中, $\boldsymbol{w}_g$ 和 b_g 为可学习参数, $g_l \in (0,1)$. 全局调节系数由当前输入特征计算, 用于对频域细化结果的贡献强度进行输入相关调节.

频域细化路径采用可学习的小波分析与重构滤波器, 并以 Haar 小波系数进行参数初始化. 滤波器参数随检测网络联合优化, 使频率分解在保留小波先验的同时适应反无人机检测任务的特征分布. 设可学习分析滤波器为 $\boldsymbol{W}_a$, 输入特征经一级离散小波变换分解为:

$$\{X^{LL}, X^{LH}, X^{HL}, X^{HH}\} = \mathrm{DWT}(X^{(l)}; \boldsymbol{W}_a). \tag{5}$$

其中, X^{LL} 为低频子带, 主要保留整体结构和上下文信息; X^{LH}、X^{HL} 和 X^{HH} 为不同方向的高频子

带, 用于表征目标边缘、轮廓和局部纹理等细节.

四个原始子带沿通道维拼接形成子带表示 Z_l, 经 1×1 卷积、批归一化和 Sigmoid 函数生成相应的重标定权重. 低频子带采用 $\widetilde{X}^{LL}=\alpha_{LL}\odot X^{LL}$ 进行直接调节, 以保持整体结构和上下文响应; 三个高频子带采用残差式重标定:

$$\widetilde{X}^{q}=X^{q}+\beta_h\alpha_q\odot X^{q},\ q\in\{LH,HL,HH\}. \quad (6)$$

其中, α_{LL} 和 α_q 分别为低频与高频子带的重标定权重, β_h 为受限可学习的高频保护系数, $\odot$ 表示逐元素乘法. 重标定权重由当前子带特征生成, 用于依据子带响应调节低频与高频分量的相对贡献. 高频子带在保留原始响应的基础上进行调节, 可避免与小尺度目标相关的细节被过度抑制.

重标定后的四个子带沿通道维重新组合, 并通过分组 3×3 卷积进行局部上下文建模. 卷积输出随后重新组织为 LL、LH、HL 和 HH 四个子带, 并利用可学习重构滤波器 $\boldsymbol{W}_r$ 完成逆离散小波变换, 得到频域细化结果 ΔX_f. AFRM 采用受限残差方式输出:

$$X_{\text{out}}=X_s+\gamma g_l\Delta X_f. \quad (7)$$

其中, $\gamma\in[0,1]$ 为受限可学习系数, 用于约束频域细化路径的整体作用强度; g_l 根据当前输入特征进一步调节频域细化结果对输出特征的贡献.

AFRM 布置于骨干网络的 L2 和 L4 特征提取位置. L2 特征具有较高空间分辨率, 有利于保持小尺度目标的边缘和纹理细节; L4 特征具有更大的感受范围, 可在进一步下采样前对目标结构与上下文响应进行调节. 两处频域细化共同作用于不同特征层级, 以保持小尺度目标的高频细节.

2.4　层级低时延特征聚合模块 (HLFA)

在固定视角可见光反无人机检测中, 颈部网络需要在不同尺度间传递目标细节与上下文信息. 常规特征聚合结构通常依赖宽通道卷积或多层卷积堆叠, 在高分辨率特征层上容易产生较大的计算与访存开销; 同时, 对不同尺度特征采用相同的处理方式, 也难以兼顾小尺度目标细节与高层语义上下文. 为此, 设计 HLFA 在 P_3、P_4 和 P_5 尺度上构建金字塔聚合路径与检测侧细化路径. 前者采用轻量递推聚合维持特征金字塔中的信息传递, 后者面向不同尺度的预测输入进行局部、区域和全局信息细化. HLFA 的整体结构如图 3 所示.

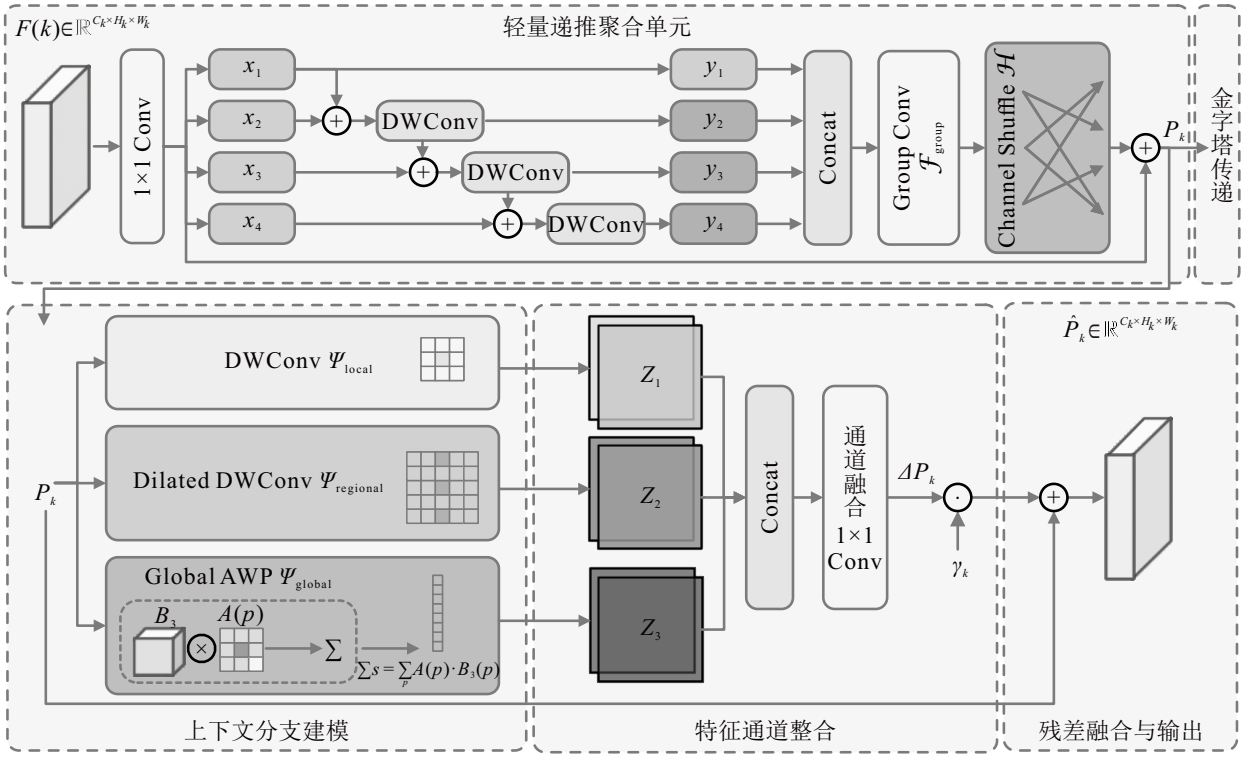

图 3　层级低时延特征聚合模块 (HLFA)

对于第 k 个尺度的输入融合特征 $F_k\in\mathbb{R}^{C_k\times H_k\times W_k}$, 其中 $k\in\{3,4,5\}$, 轻量递推聚合单元首先通过 1×1 卷积完成通道投影, 并沿通道维划分为 S 个通道子组 $\{x_1,x_2,\ldots,x_S\}$. 各通道子组采用逐级递推方式进行信息聚合:

$$y_i=\begin{cases}x_1, & i=1\\ \mathcal{D}_i(x_i+y_{i-1}), & 2\leqslant i\leqslant S\end{cases} \quad (8)$$

其中, $\mathcal{D}_i(\cdot)$ 表示作用于第 i 个通道子组的 3×3 深

度卷积. 第一个通道子组保持直接传递, 其余通道子组在融合前一组输出后进行局部特征提取, 使层级信息能够在不同通道分组之间逐步传递. 尽管递推过程保留了相邻通道子组间的信息依赖, 但各卷积仅作用于较窄的通道分组, 从而避免在完整通道维度上重复执行空间卷积, 降低特征聚合过程中的计算与访存开销.

各通道子组的输出沿通道维拼接后, 经分组 1×1 卷积和通道重排完成轻量通道融合, 以促进跨组信息交互; 随后与输入映射残差相加, 得到当前尺度的金字塔聚合特征 P_k. 聚合特征 P_k 沿原有特征金字塔路径继续传递, 并作为当前尺度检测侧细化路径的输入.

检测侧细化路径与金字塔聚合路径同层设置. 该路径将聚合特征 P_k 同时送入局部、区域和全局三个并行分支. 局部分支采用 3×3 深度卷积提取目标边缘与纹理信息; 区域分支采用空洞 3×3 深度卷积建模邻域上下文; 全局分支采用自适应加权池化 (Adaptive Weighted Pooling, AWP), 根据空间响应生成归一化权重并汇聚全局上下文, 随后通过通道映射形成全局校准响应. 三个分支的输出经受限门控调节后分别记为 Z_1、Z_2 和 Z_3, 随后沿通道维拼接, 并通过通道融合生成第 k 个尺度的细化增量 ΔP_k.

不同尺度采用具有差异化侧重的门控配置: P_3 侧重局部空间细节, 以保持小尺度目标的边缘与纹理响应; P_4 侧重区域上下文, 以协调目标结构与邻域信息; P_5 侧重全局上下文, 以增强高层语义表达.

检测侧细化特征采用受限残差方式输出:

$$\widehat{P}_k = P_k + \gamma_k \Delta P_k,\ k\in\{3,4,5\} \tag{9}$$

其中, γ_k 为第 k 个尺度对应的受限可学习残差系数, 用于控制检测侧细化增量的整体贡献强度. 聚合特征 P_k 沿原有特征金字塔路径继续传递, 检测侧细化特征 $\widehat{P}_k$ 则直接送入对应尺度的预测层.

3 实 验

3.1 数据集与实验设置

为验证所提方法在复杂背景条件下的无人机目标检测性能, 本节采用 DUT Anti-UAV [16] 数据集开展实验. 该数据集包含约 10,000 张可见光图像样本, 覆盖复杂背景干扰、姿态变化及非均匀光照等典型场景因素, 适用于评估固定视角下的无人机目标检测能力. 该数据集样本均为可见光图像, 不包含夜间红外场景图像, 因此本节实验范围限定于固定视角可见光反无人机检测任务.

实验部分包括消融实验、对比实验、边缘目标图像组检测性能分析、运行效率分析、视场边界风险联合分析和可视化分析. 其中, 消融实验用于分析 AFRM 和 HLFA 等关键结构的贡献; 对比实验用于评估所提方法相对于主流检测模型的综合性能; 边缘目标图像组检测性能分析用于评价模型在边缘目标图像组上的整体检测效果; 运行效率分析用于统计模型在统一硬件平台上的推理时延与帧率; 视场边界风险联合分析进一步利用检测输出评估边缘目标形成稳定观测输入的风险; 可视化分析用于展示不同场景下所提方法与基线模型在目标检出状态、预测置信度和边缘目标风险状态上的差异.

考虑到远距离无人机在图像中通常呈现为小尺度目标, 实验将归一化标注框统一映射至 640×640 参考坐标, 并按照目标最大边长 $l=\max(w,h)$ 进行尺度统计. 具体地, 将目标划分为极小目标 ($l\leqslant16$ px)、小目标 ($16<l\leqslant40$ px)、中目标 ($40<l\leqslant96$ px) 和大目标 ($l>96$ px) 四个连续区间, 用于分析数据集中的目标尺度分布. 如图 4 所示, 各柱从下到上依次表示极小、小、中和大目标; DUT Anti-UAV 数据集包含较高比例的小尺度目标, 测试集在四个尺度区间上的分布相对更均衡, 为分析模型在小尺度目标及不同目标尺度条件下的检测表现提供了样本基础.

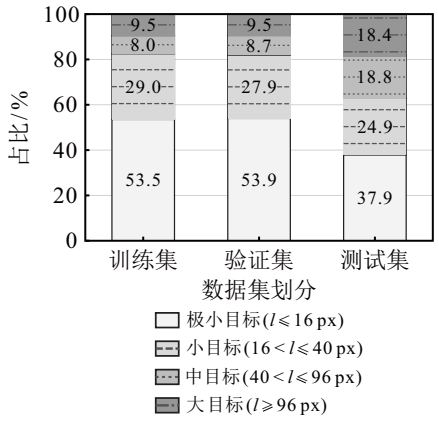

图 4 DUT Anti-UAV 数据集目标尺度分布

为分析图像边界附近目标的检测表现, 实验沿用前述视场边界余量 $d_e=\min(c_x,W-c_x,c_y,H-c_y)$ 筛选边缘真实目标. 当 $d_e\leqslant\alpha\min(W,H)$ 时, 该目标计为边缘真实目标. 实验在模型预测结果统计前预先设定 $\alpha=0.20$, 边缘真实目标的筛选仅依赖测试集真实标注框的位置, 不使用任何模型的检测结果. 该阈值对应于图像四周宽度为图像短边 20% 的边界带区域. 基于该定义, 包含至少一个边缘真实目标的测试图像构成边缘目标图像组, 用于评

价模型在边缘目标相关图像上的整体检测效果; 边缘真实目标集合用于视场边界风险联合分析.

所有实验在配备 Intel Xeon Platinum 8352V CPU 和 NVIDIA GeForce RTX 4090 GPU(24 GB) 的硬件平台上进行. 软件环境为 Ubuntu 20.04.1、CUDA 12.4 和 PyTorch 2.5.1. 所有模型均未采用预训练权重, 从头训练. 模型输入尺寸为 640 × 640, 训练 300 轮, 批大小设置为 32, 随机种子设置为 42. 对于适用模型, 初始学习率、动量系数和权重衰减分别设置为 0.01、0.937 和 0.0005; 马赛克增强概率设置为 1.0, 并在最后 10 轮关闭; 测试阶段的 NMS IoU 阈值设置为 0.7. 除模型结构及其特定优化与后处理设置外, 各对比模型与消融模型采用一致的数据划分、训练轮数、批大小和测试数据.

3.2　评价指标

在检测性能评估方面, 实验采用精确率 (Precision, P) 和召回率 (Recall, R) 分别衡量检测结果的准确性和目标检出完整性, 并采用 Precision–Recall 曲线下的平均精度 (Average Precision, AP) 作为主要评价指标. 实验报告 AP_{50} 和 $AP_{50:95}$, 其中, AP_{50} 表示 IoU 阈值为 0.50 时的平均精度, $AP_{50:95}$ 表示 IoU 阈值从 0.50 增加至 0.95、步长为 0.05 时所得 AP 的均值, 用于综合评估模型的检测与定位性能.

在模型复杂度方面, 实验采用参数量 (Params) 和浮点运算量 (FLOPs) 衡量模型规模与理论计算开销. Params 反映模型的可学习参数规模, FLOPs 反映模型在 640 × 640 输入分辨率下完成单次前向传播所需的理论计算量.

在运行效率评估方面, 实验采用平均单帧模型执行时延 (ms/image)、90% 分位单帧模型执行时延 (P90) 和帧率 (FPS) 衡量模型执行效率. 其中, FPS 由平均单帧模型执行时延换算得到. 所有运行效率指标均在批大小为 1 的设置下统计, 以反映单帧输入条件下检测模型的执行效率.

在边缘目标图像组检测性能分析中, 实验仍采用 Precision、Recall、AP_{50} 和 $AP_{50:95}$ 作为标准检测指标. 该评价以含边缘真实目标的测试图像为分析对象, 用于评价模型在边缘目标图像组上的整体检测效果.

在视场边界风险联合分析中, 实验结合 FBRR、检测状态和预测置信度计算单目标级视场边界联合风险指数 R_i, 并对边缘真实目标进行等权平均, 得到平均联合风险指数 $\bar{R}$. 上述指标数值越低, 表示模型在本文统一实验条件下对应的边缘目标联合风险越低, 具体定义与计算方法见后文.

3.3　消融实验

为验证 AFRM、HLFA 及其关键机制的有效性, 本节以 YOLO26n 为基线模型开展消融实验. 实验包括模块级与机制级两个层面: 模块级实验考察完整 AFRM、完整 HLFA 及二者组合后的性能; 机制级实验进一步分析 AFRM 中频率分解、子带重标定和高频残差保护机制, 以及 HLFA 中分组递推、尺度相关分支门控和双路径特征传递方式的作用. 消融结果如表 1 所示.

表 1　模块及关键机制消融实验结果

Model	Precision	Recall	AP_{50}	$AP_{50:95}$	Params (M)	FLOPs (G)
Base (YOLO26n)	0.962	0.904	0.950	0.662	2.5	5.8
AFRM(仅固定 Haar 变换)	0.954	0.905	0.942	0.651	2.2	5.6
AFRM(无子带重标定)	0.960	0.922	0.955	0.663	2.3	5.6
AFRM(高频子带直接重标定)	0.961	0.917	0.956	0.668	2.3	5.6
AFRM	0.967	0.915	0.962	0.672	2.2	5.6
HLFA(无分组递推)	0.955	0.908	0.952	0.661	2.5	5.5
HLFA(统一分支门控配置)	0.968	0.909	0.952	0.662	2.2	5.4
HLFA(细化特征参与金字塔传递)	0.963	0.907	0.951	0.648	2.2	5.5
HLFA	0.968	0.917	0.957	0.667	2.2	5.4
FLA-Net (AFRM + HLFA)	0.969	0.914	0.960	0.678	2.2	5.4

在 AFRM 的内部机制消融中, 仅采用固定 Haar 变换时, AP_{50} 和 $AP_{50:95}$ 分别为 0.942 和 0.651. 固定 Haar 小波对不同输入采用一致的预设频带划分, 在目标高频细节与复杂背景纹理同时存在时, 其特征重组方式与检测任务的优化目标可能存在一定适配偏差, 因此该设置的指标低于基线模型. 完整 AFRM 以 Haar 系数作为初始化, 并通过可学习滤波器和子带重标定对频率响应进行任务相关调整, 取得了更好的检测结果, 表明性能收益来自频域先验与自适应调节机制的协同作用. 采用统一的子带响应进行特征重构时, 模型取得了较高的 Recall, 而完整 AFRM 在 Precision、AP_{50} 和 $AP_{50:95}$ 上表现更

优, 说明子带重标定有助于将较宽的目标响应进一步聚焦到更具判别性的频率成分. 高频子带采用直接重标定时, 两项 AP 分别为 0.956 和 0.668; 残差式重标定则在原始高频响应上叠加受控增量, 使目标边缘与纹理信息保持稳定的传递路径, 其结果相较直接重标定进一步提高, 与高频残差保护机制的设计目的保持一致.

在 HLFA 的内部机制消融中, 采用独立通道分组处理时, $AP_{50:95}$ 为 0.661; 完整 HLFA 通过相邻子组间的逐级递推建立层级信息联系, 使局部响应能够在较窄通道范围内逐步聚合, 对应的 $AP_{50:95}$ 提高至 0.667. 这一结果表明, 分组递推能够在维持轻量计算方式的同时促进子组间的信息交互. 检测侧细化路径进一步根据不同尺度的特征特点设置分支门控偏置: P_3 侧重局部分支, 以保持目标边缘与纹理细节; P_4 侧重区域分支, 以协调目标结构与邻域上下文; P_5 侧重全局分支, 以增强高层语义校准, 各分支门控均在受限区间内随网络联合优化. 采用统一分支门控配置时, AP_{50} 和 $AP_{50:95}$ 分别为 0.952 和 0.662; 完整 HLFA 则达到 0.957 和 0.667, 表明尺度相关门控配置能够更好地适应不同检测层级的特征表达需求. 对于双路径传递方式, 完整 HLFA 使用聚合特征维持金字塔中的跨尺度传递, 并将细化特征直接送入相应预测层, 使跨尺度信息聚合和尺度相关预测细化分别承担明确功能. 相较细化特征参与金字塔传递的设置, 该方式取得了更高的综合检测指标, 说明这种路径分工更适合当前多尺度无人机目标检测任务.

从模块级结果来看, 基线模型的 AP_{50} 和 $AP_{50:95}$ 分别为 0.950 和 0.662. 完整 AFRM 将两项指标提高至 0.962 和 0.672, 完整 HLFA 则在 Params 和 FLOPs 降至 2.2 M 和 5.4 G 的同时, 取得了 0.957 和 0.667 的 AP_{50} 与 $AP_{50:95}$. 组合两个模块后, FLA-Net 的 Precision 和 $AP_{50:95}$ 分别达到 0.969 和 0.678. 与完整 AFRM 相比, FLA-Net 的 Recall 和 AP_{50} 变化较小, 而 $AP_{50:95}$ 提高了 0.006, 说明两个模块组合后保持了较好的目标检出覆盖, 并在更严格的 IoU 阈值范围内获得了更优的综合定位表现. 该结果与 AFRM 侧重频域细节调节、HLFA 侧重多尺度聚合和预测侧细化的功能定位相符.

3.4 对比实验

为验证 FLA-Net 在固定视角反无人机检测任务中的综合性能, 本节在 DUT Anti-UAV 测试集上与多类代表性检测模型进行对比. 对比方法包括反无人机检测方法 YOLOv8-DAW [13], 以及 TOOD [17]、RTMDet-Tiny [18]、DINO [19]、Faster R-CNN [20]、Cascade R-CNN [21] 和 RT-DETR [22] 等代表性检测模型; 此外, 还选取轻量 YOLO 系列模型 YOLO8n [23]、YOLO9t[24]、YOLO10n[25]、YOLO11n[26]、YOLO12n[27]、YOLO13n [28] 和 YOLO26n [29] 进行对比. 所有模型均采用 DUT Anti-UAV 官方测试集进行评价, 评价指标包括 Precision、Recall、AP_{50}、$AP_{50:95}$、Params 和 FLOPs. 固定视角反无人机检测需要兼顾检测性能、定位质量和计算开销, 因此从上述三个方面对不同方法进行综合比较, 结果如表 2 所示.

表 2 对比实验结果

模型	Precision	Recall	AP_{50}	$AP_{50:95}$	Params (M)	FLOPs (G)
YOLOv8-DAW[13]	0.947	0.834	0.890	0.571	2.1	–
TOOD[17]	0.835	0.817	0.842	0.507	32.0	186.0
RTMDet-Tiny[18]	0.950	0.883	0.929	0.643	4.9	8.0
DINO[19]	**0.976**	0.950	**0.963**	0.667	47.5	273.6
Faster R-CNN[20]	0.954	0.905	0.933	0.632	41.3	207.9
Cascade R-CNN[21]	0.965	0.909	0.945	0.616	69.2	235.7
RT-DETR[22]	0.843	**0.961**	0.925	0.624	42.9	134.8
YOLO8n[23]	0.959	0.910	0.945	0.650	3.0	8.1
YOLO9t[24]	0.964	0.896	0.938	0.638	**2.0**	7.6
YOLO10n[25]	0.941	0.906	0.937	0.641	2.7	8.2
YOLO11n[26]	0.953	0.903	0.936	0.640	2.6	6.4
YOLO12n[27]	0.961	0.903	0.937	0.635	2.6	6.5
YOLO13n[28]	0.960	0.901	0.934	0.635	2.5	6.4
YOLO26n[29]	0.962	0.904	0.950	0.662	2.5	5.8
FLA-Net (Ours)	0.969	0.914	0.960	**0.678**	2.2	**5.4**

从表 2 可以看出, 不同类型的检测模型在各项指标上呈现出不同特点. 在当前实验设置下, DINO 的 Precision 和 AP_{50} 分别达到 0.976 和 0.963, 均为表中最高; RT-DETR 的 Recall 达到 0.961, 在目标检出覆盖方面表现突出. FLA-Net 的 Precision 和 AP_{50} 分别为 0.969 和 0.960, 均取得次优结果, $AP_{50:95}$ 则达到表中最高的 0.678, 表明其在保持较好目标判别与检出性能的同时, 在多个 IoU 阈值下具有较好的综合定位表现. FLA-Net 的参数量为 2.2 M, 接近表中参数量最小的 YOLO9t; 其 FLOPs 为 5.4 G, 在已统计 FLOPs 的对比模型中最低. 相较直接基线 YOLO26n, FLA-Net 在提高检测性能的同时降低了参数量和 FLOPs, 并在检测质量与计算开销之间取得了较好的平衡.

从模型类型来看, 轻量 YOLO 系列模型和 RTMDet-Tiny 的模型复杂度整体较低, 而 DINO、Faster R-CNN、Cascade R-CNN 和 RT-DETR 的参数

量与计算量较大. FLA-Net 以 2.2 M 参数量和 5.4 G FLOPs 取得最高的 $AP_{50:95}$, 体现了综合定位性能与模型复杂度之间的平衡.

综上, 对比实验表明, FLA-Net 在 DUT Anti-UAV 测试集上取得了较好的检测性能与模型复杂度平衡. 表 2 给出了测试集整体检测性能, 后续进一步构建边缘目标图像组, 对含边缘真实目标图像上的标准检测性能进行补充分析.

3.5　边缘目标图像组检测性能分析

整体测试集评价能够反映模型的平均检测能力, 而靠近视场边界的无人机目标具有较小的边界余量, 其可见区域和目标框位置更容易受到图像边界约束. 为进一步分析模型在边缘目标相关样本上的标准检测性能, 本节依据前文定义的边缘真实目标判定准则, 并取 $\alpha = 0.20$, 筛选测试集中包含至少一个边缘真实目标的图像, 构成边缘目标图像组. 该图像组共包含 278 张测试图像, 其中边缘真实目标数量为 280 个, 图像组内真实目标总数为 296 个, 边缘真实目标占比为 $280/296 \approx 94.6\%$. 本节标准检测指标以图像组内全部 296 个真实目标为评价对象, 检测结果如表 3 所示.

表 3　边缘目标图像组检测性能对比

模型	Precision	Recall	AP_{50}	$AP_{50:95}$
YOLO11n[26]	**0.991**	0.916	0.951	0.701
YOLO12n[27]	0.968	0.912	0.946	0.690
YOLO26n[29]	0.966	0.929	0.954	0.709
FLA-Net	0.973	**0.932**	**0.967**	**0.723**

从表 3 可以看出, YOLO11n 的 Precision 达到 0.991, 为表中最高. FLA-Net 的 Precision 为 0.973, 同时在 Recall、AP_{50} 和 $AP_{50:95}$ 上分别取得 0.932、0.967 和 0.723, 均为表中最高结果. 与直接基线 YOLO26n 相比, FLA-Net 的 Recall、AP_{50} 和 $AP_{50:95}$ 分别提高了 0.003、0.013 和 0.014, 表明其在保持较高精确率的同时, 具有更好的目标检出覆盖和综合定位表现. 考虑到边缘真实目标占该图像组全部真实目标的 94.6%, 上述结果主要体现了模型在边缘目标相关图像上的检测与定位性能.

综上, 边缘目标图像组实验表明, FLA-Net 在边缘目标相关样本上具有较好的检出表现和定位质量. 该实验完成了图像组层面的标准检测性能评价, 后续进一步引入模型运行时延, 并结合边缘目标的边界余量、检测状态和预测置信度, 开展目标级视场边界风险联合分析.

3.6　运行效率分析

固定视角反无人机检测任务通常面向连续图像输入, 要求检测模型具备较低且稳定的单帧模型执行时延. 相比训练框架内的 PyTorch 前向测试, ONNX 模型能够在独立推理后端下评估网络结构的模型执行效率, 更接近面向推理应用的模型运行形态. 基于此, 本节采用 ONNX Runtime 后端对不同轻量检测模型进行运行效率测试, 统计其模型执行时延与帧率.

具体而言, 各模型的 PyTorch 权重首先被导出为固定输入尺寸 640×640 的 ONNX 模型. 随后, 在 RTX 4090 实验平台上统一采用 ONNX Runtime CUDA Execution Provider, 以 batch size=1 对完整 DUT Anti-UAV 测试集图像进行逐图推理测速. 完整测试集能够提供较多的计时样本, 从而减小少量样本测速造成的偶然波动. 考虑到本节主要关注轻量实时检测模型, 选取 YOLO11n [26]、YOLO12n [27] 和 YOLO26n [29] 作为对比对象, 其中 YOLO26n 为所提方法的直接基线.

测速前, 所有测试图像均完成预处理并缓存为模型输入, 正式计时阶段不计入磁盘 I/O 和图像预处理开销. 每个模型在正式测速前执行 50 次预热推理, 以减小模型加载、算子初始化和缓存建立造成的时间波动. 正式测速阶段中, 每个模型重复测试 5 轮, 每轮遍历完整 DUT Anti-UAV 测试集, 并记录逐图 ONNX Runtime 模型执行时延. 最终统计平均单帧时延 Mean(ms)、90% 分位时延 P90(ms)、帧率 FPS 以及 5 轮平均时延的标准差 Std(ms), 结果如表 4 所示.

表 4　ONNX Runtime 后端运行效率对比

模型	Mean (ms)	P90 (ms)	FPS	Std (ms)
YOLO11n[26]	5.15	6.02	194.01	0.26
YOLO12n[27]	6.20	6.90	161.28	**0.14**
YOLO26n[29]	5.30	6.11	188.65	0.19
FLA-Net	**4.94**	**5.64**	**202.59**	0.26

从表 4 可以看出, FLA-Net 的平均单帧时延为 4.94 ms, 对应帧率为 202.59 FPS, 均优于对比模型, 表明其在 ONNX Runtime 后端下具有较高的平均模型执行效率. FLA-Net 的 P90 时延为 5.64 ms, 同样为表中最低, 即 90% 的逐图模型执行时延不超过该数值. 其 5 轮平均时延的标准差为 0.26 ms, 与 YOLO11n 相当. 综合平均时延、P90 时延和轮间波动可以看出, FLA-Net 在平均响应速度和高分位时延方面均具有较好表现.

综上, 运行效率分析表明, FLA-Net 在完整 DUT Anti-UAV 测试集上取得了较低的 ONNX Runtime 模型执行时延和较高帧率. 该结果为后续视场边界风险联合分析提供了统一推理后端下的时延参数; 结合前述边缘目标图像组检测性能, 可进一步分析检测状态、预测置信度、边界余量和模型响应时间共同作用下的边缘目标观测风险.

3.7 基于 FBRR 的视场边界风险联合分析

在固定视角反无人机检测任务中, 靠近图像边界的目标具有更有限的视场边界余量. 在局部匀速并朝向最近视场边界运动的保守假设下, 目标在单帧推理期间的估计图像位移将消耗其当前边界余量, 从而缩短目标保持在视场内的估计时间窗口. 为评估边缘目标在检测链路中的边界观测风险, 本节基于边缘真实目标集合开展视场边界风险联合分析, 并结合视场边界风险比 (field-of-view boundary risk ratio, FBRR)、检测状态和预测置信度计算单目标级视场边界联合风险指数 R_i. 联合分析结果采用 N_e(边缘真实目标数量)、TP 数量、FN 数量、$N_{\rho\geqslant 1}$(边界风险饱和目标数量) 和 $\bar{R}$(平均联合风险指数) 进行汇总, 实验结果如表 5 所示.

表 5 边缘真实目标集合的视场边界风险联合分析结果

模型	τ(ms)	N_e	TP	FN	$N_{\rho\geqslant 1}$	$\bar{R}$
YOLO11n[26]	5.15	280	259	21	22	0.50
YOLO12n[27]	6.20	280	255	25	28	0.55
YOLO26n[29]	5.30	280	261	19	23	0.48
FLA-Net	**4.94**	280	**263**	**17**	22	**0.47**

注: τ 为表4 中 ONNX Runtime 后端平均单帧时延.

本节沿用前文边缘真实目标定义, 并取 $\alpha = 0.20$ 筛选测试集中的边缘真实目标. 该阈值对应于图像四周宽度为短边 20% 的边缘区域, 与边缘目标图像组检测性能分析保持一致. 不同的是, 本节以边缘真实目标为基本统计单位, 用于开展目标级视场边界风险分析.

实验输入包括 DUT Anti-UAV 测试集标注、各模型预测结果以及单帧模型执行时延 τ. 其中, τ 采用运行效率分析中 ONNX Runtime 后端的平均单帧模型执行时延, 并换算为秒参与 FBRR 计算, 用于在统一平均运行条件下比较不同检测器的视场边界风险. 测试集标注用于确定边缘真实目标及其边界余量, 模型预测结果用于获得边缘真实目标的检测状态和 TP 目标对应的预测置信度. 风险分析仅以边缘真实目标为统计对象, 预测框经置信度阈值 0.25 筛选后, 在 IoU 阈值为 0.50 的条件下确定 TP/FN 状态. 需要说明的是, 边缘目标图像组检测性能以 278 张图像内的全部 296 个真实目标为评价对象, 而本节仅统计其中的 280 个边缘真实目标, 因此两部分 Recall 的统计范围不同.

根据 DJI 官方规格[30], 主流民用四旋翼无人机如 Mini 4 Pro、Air 3 和 Mavic 3 Pro 的最高水平速度分别为 16、21 和 21 m/s, 展开状态最大外形跨度分别约为 0.373、0.326 和 0.348 m; 据此, 本文取代表性物理速度 $v = 20\,\mathrm{m/s}$, 并以代表性机体跨度 $L = 0.35\,\mathrm{m}$ 作为图像尺度换算的物理基准. 对于第 i 个边缘真实目标, 记其图像尺度为 $l_i = \max(w_i, h_i)$, 边界余量为 $d_{e,i}$, 则单帧模型执行时延内的估计图像位移量 Δp_i 和 FBRR ρ_i 分别为

$$\Delta p_i = \frac{v\tau l_i}{L},\ \rho_i = \frac{\Delta p_i}{d_{e,i}}. \tag{10}$$

其中, τ 在计算时换算为秒, Δp_i 表示由物理运动速度和单帧推理时延换算得到的图像平面位移估计, ρ_i 表示该位移相对于当前视场边界余量的比例.

在获得 ρ_i 后, 结合检测状态和预测置信度计算单目标级视场边界联合风险指数. 对于 TP 目标, 记其预测置信度为 c_i. 当 $\rho_i \geqslant 1$ 时, 估计图像位移达到或超过当前视场边界余量, 表示目标在单帧推理时延内可能接近或越过视场边界. 因此, 采用 $\min(\rho_i, 1)$ 作为饱和后的边界风险项, 并将 R_i 定义为

$$R_i = \begin{cases} 1, & \text{FN}, \\ 1 - (1 - \min(\rho_i, 1))c_i, & \text{TP}. \end{cases} \tag{11}$$

其中, FN 目标的 R_i 取最大值; TP 目标的 R_i 随 ρ_i 增大而增大, 随预测置信度提高而降低.

最后, 为保持目标级统计口径一致, $\bar{R}$ 采用边缘真实目标等权平均计算, 用于表征模型在本文统一物理参数、目标匹配协议和运行时延条件下的平均联合风险. FN 目标以最大风险计入, TP 目标的风险由 FBRR 和预测置信度共同决定; $\bar{R}$ 数值越低, 表示模型在上述统一实验条件下对应的平均联合风险越低. $N_{\rho\geqslant 1}$ 表示 FBRR 达到或超过 1 的边缘真实目标数量, 用于反映边界风险进入饱和区间的目标规模. 边缘区域虚警属于预测框级风险, 可作为后续扩展方向进一步纳入统一的边缘视场风险建模.

从表 5 可以看出, FLA-Net 在边缘真实目标集合上取得最高 TP 数量和最低 FN 数量, 表明在当前匹配协议下, 更多边缘真实目标能够与模型预测结果形成有效匹配. 同时, FLA-Net 的平均单帧模型执行时延最低, 使目标在单帧模型执行时间内产生的

估计图像位移相对更小. 在综合检测状态、预测置信度、边界余量和推理时延后, FLA-Net 的 $\bar{R}$ 最低, 表明其在本文指标定义与实验参数下具有较低的平均联合风险.

进一步比较不同模型可以看出, YOLO12n 的 FN 数量和 $N_{\rho\geqslant 1}$ 均较高, 对应 $\bar{R}$ 也最高, 表明该模型在当前检测状态与时延条件下对应的平均联合风险较高. YOLO11n 的平均单帧推理时延仅次于 FLA-Net, 但综合其 TP/FN 状态、预测置信度和边界余量后, 平均联合风险指数仍高于 FLA-Net, 说明模型执行时延与检测响应需要在联合指标中共同考虑. YOLO26n 的 TP/FN 表现接近 FLA-Net, 在进一步结合预测置信度、边界余量和模型执行时延后, 其平均联合风险指数为 0.48, 略高于 FLA-Net 的 0.47. 上述结果说明, 视场边界风险联合分析能够在标准检测性能之外, 进一步反映边缘目标在有限视场余量和检测响应时间约束下的风险差异. 下一节将结合典型样例, 对不同风险条件下的检测结果进行定性分析.

3.8　可视化分析

为进一步从定性角度分析边缘目标检测结果与视场边界联合风险之间的关系, 本节从边缘真实目标集合中选取四类典型样例进行可视化对比, 如图 5 所示. 图中每列对应一种边缘目标情形, 依次为低 FBRR 条件下的检测响应增强、小尺度边缘目标、低 FBRR 漏检恢复和高 FBRR 漏检恢复; 各行依次为真实标注、YOLO26n 基线模型和 FLA-Net. 为提高图中信息可读性, 真实标注行的相应子图上方给出边界余量和目标尺度, 两个模型检测结果上方给出检测状态、预测置信度和单目标联合风险 R_i. 其中, FBRR 表示单帧推理期间估计图像位移与当前视场边界余量的比例, 用于综合反映目标尺度、边界余量和模型执行时延对应的边界风险状态.

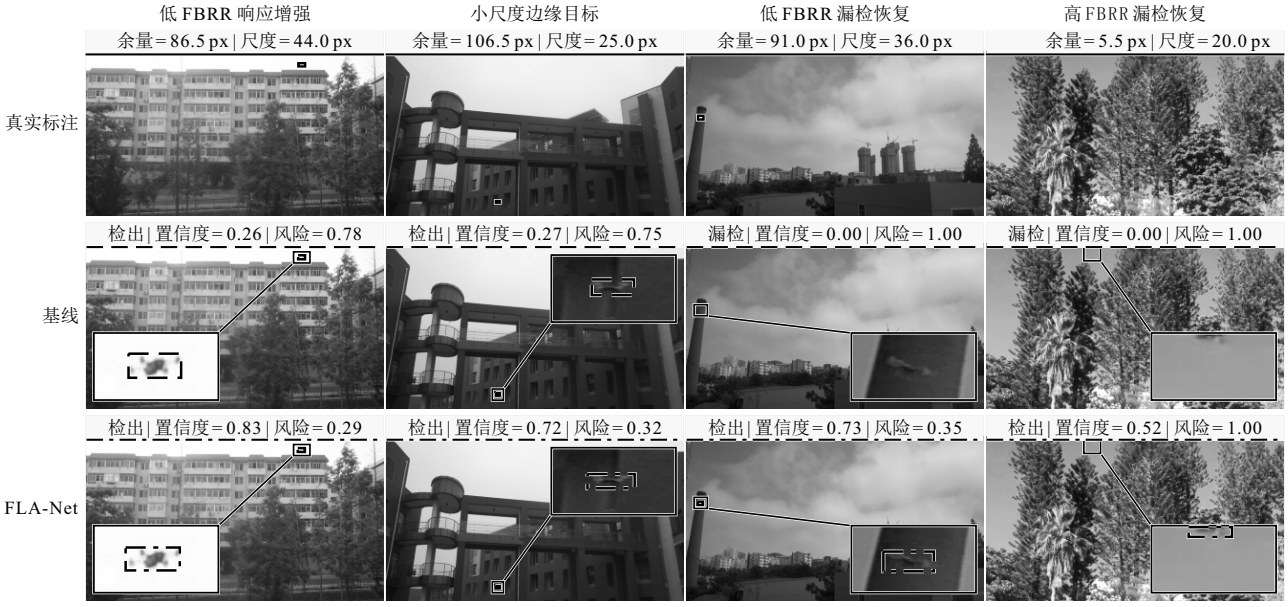

图 5　边缘目标检测结果与视场边界联合风险可视化对比

上述四类样例分别从预测响应、目标尺度、检出状态和边界余量四个方面补充说明前述定量结果.

从图 5 可以看出, 在前两类样例中, 基线模型与所提方法均能够给出有效检测结果, 所提方法具有更高的预测置信度和更低的单目标联合风险. 在低 FBRR 漏检恢复样例中, 基线模型未能检出边缘目标, 其 R_i 取最大值; 所提方法检出目标后, 按照本文联合风险定义计算得到的 R_i 明显降低. 在高 FBRR 漏检恢复样例中, 所提方法虽然成功检出目标, 但由于目标边界余量较小, FBRR 已达到或超过 1, 边界风险项进入饱和区间, 因此联合风险仍保持较高水平. 上述结果表明, FLA-Net 能够提高边缘目标的检出能力与预测置信度, 并在低 FBRR 条件下降低单目标联合风险; 在高 FBRR 条件下, 所提方法仍能实现对基线漏检目标的有效检出, 体现了其在有限边界余量场景中的检测优势.

4　结　论

针对固定视角反无人机检测中远距离边缘小尺度目标在有限视场边界余量下难以形成稳定观测输入的问题, 设计了视场边界风险比 FBRR 和单目标联合风险 R_i, 将边界余量、目标尺度、检测状态、预测置信度和单帧模型执行时延纳入统一分析. 基于该场景约束, 提出频域-时延协同优化检测框架 FLA-Net, 通过 AFRM 与 HLFA 增强边缘小尺度目标的细粒度特征表达和跨尺度信息聚合能力, 并在

较低计算开销下兼顾低时延推理. 在 DUT Anti-UAV 测试集上的实验结果表明, FLA-Net 在整体检测性能、边缘目标图像组检测性能和 ONNX Runtime 后端运行效率方面取得了较优的综合表现; 视场边界风险联合分析进一步表明, 所提方法能够为边缘目标提供更稳定的检测输出, 并降低由漏检、低置信度响应和有限边界余量共同导致的观测风险. 未来工作将面向夜间红外和复杂背景干扰等场景验证模型泛化能力, 并将边缘区域虚警、连续帧运动状态纳入统一风险分析框架.

参考文献 (References)

[1] 邓余婉祺, 王越, 杨超, 等. 无人机语义安全研究综述[J]. 计算机学报, 2025, 48(6): 1495-1515.
(Deng Y W Q, Wang Y, Yang C, et al. A survey on semantic safety of unmanned aerial vehicle[J]. Chinese Journal of Computers, 2025, 48(6): 1495-1515.)

[2] Dong Y F, Wu F Y, Zhang S J, et al. Securing the skies: A comprehensive survey on anti-UAV methods, benchmarking, and future directions[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville, 2025: 6725-6739.

[3] Ahmad B I, Rogers C, Harman S, et al. A review of automatic classification of drones using radar: Key considerations, performance evaluation, and prospects[J]. IEEE Aerospace and Electronic Systems Magazine, 2024, 39(2): 18-33.

[4] 交通运输部. 民用无人驾驶航空器运行安全管理规则: 交通运输部令 2024 年第 1 号[EB/OL]. (2024-01-03)[2026-08-06]. https://xxgk.mot.gov.cn/2020/jigou/fgs/202401/t20240103_3980642.html.
(Ministry of Transport of the People's Republic of China. Rules for the safety management of civil unmanned aircraft operations: Order No. 1 of 2024[EB/OL]. (2024-01-03)[2026-08-06]. https://xxgk.mot.gov.cn/2020/jigou/fgs/202401/t20240103_3980642.html.)

[5] Ding G H, Ren Y H, Liu Y T, et al. Vision-based anti-unmanned aerial technology: Opportunities and challenges[J]. IEEE Geoscience and Remote Sensing Magazine, 2025, 13(4): 382-405.

[6] 李沁檀, 康雪玲, 陈伟文, 等. 基于自身特性的反无人机目标检测技术研究进展[J]. 航空学报, DOI: 10.7527/S1000-6893.2026.33268.
(Li Q T, Kang X L, Chen W W, et al. Research Progress on anti-UAV target detection[J]. Acta Aeronautica et Astronautica Sinica, DOI: 10.7527/S1000-6893.2026.33268.)

[7] 王子玉, 张世勇, 董千里, 等. 面向目标实时稳定跟踪的无人机主动观测轨迹规划[J]. 控制与决策, 2025, 40(11): 3220-3228.
(Wang Z Y, Zhang S Y, Dong Q L, et al. Active observation trajectory planning for real-time and stable target tracking with UAVs[J]. Control and Decision, 2025, 40(11): 3220-3228.)

[8] Al-lQubaydhi N, Alenezi A, Alanazi T, et al. Deep learning for unmanned aerial vehicles detection: A review[J]. Computer Science Review, 2024, 51: 100614.

[9] Wang G, Yang X, Li L, et al. Tiny drone object detection in videos guided by the bio-inspired magnocellular computation model[J]. Applied Soft Computing, 2024, 163: 111892.

[10] Han J, Cao R, Brighente A, et al. Light-YOLOv5: A lightweight drone detector for resource-constrained cameras[J]. IEEE Internet of Things Journal, 2024, 11(6): 11046-11057.

[11] Zhang R Q, Shao Z F, Huang X, et al. Adaptive dense pyramid network for object detection in UAV imagery[J]. Neurocomputing, 2022, 489: 377-389.

[12] 张开玉, 王浩杰, 卢迪. FEMS-YOLO: 基于特征增强多尺度融合的无人航拍图像目标检测[J]. 控制与决策, 2026, 41(8): 2265-2275.
(Zhang K Y, Wang H J, Lu D. FEMS-YOLO: An UAV image target detection based on feature-enhanced multi-scale fusion[J]. Control and Decision, 2026, 41(8): 2265-2275.)

[13] 刘思成, 高伟伟, 侯长波, 等. YOLOv8-DAW: 一种反低空无人机目标检测算法[J]. 电光与控制, 2026, 33(4): 22-28.
(Liu S C, Gao W W, Hou C B, et al. YOLOv8-DAW: An object detection algorithm against low-altitude UAVs[J]. Electronics Optics & Control, 2026, 33(4): 22-28.)

[14] Yang Y T, Jiao L C, Liu X, et al. Dual wavelet attention networks for image classification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(4): 1899-1910.

[15] Wang Y F, Zhao J, Fan Z X, et al. JTD-UAV: MLLM-enhanced joint tracking and description framework for anti-UAV systems[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, 2025: 1633-1644.

[16] Zhao J, Zhang J S, Li D D, et al. Vision-based anti-UAV detection and tracking[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 25323-25334.

[17] Feng C J, Zhong Y J, Gao Y, et al. TOOD: Task-aligned one-stage object detection[C]. IEEE/CVF International Conference on Computer Vision. Montreal, 2021: 3490-3499.

[18] Lyu C, Zhang W, Huang H, et al. RTMDet: An empirical study of designing real-time object detectors[PP/OL]. 2022, arXiv: 2212.07784.

[19] Zhang H, Li F, Liu S, et al. DINO: DETR with improved denoising anchor boxes for end-to-end object detection[C]. Proceedings of the International Conference on Learning Representations. Kigali, 2023.

[20] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal

networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[21] Cai Z W, Vasconcelos N. Cascade R-CNN: High quality object detection and instance segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(5): 1483-1498.

[22] Zhao Y A, Lv W Y, Xu S L, et al. DETRs beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2024: 16965-16974.

[23] Varghese R, Sambath M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]. International Conference on Advances in Data Engineering and Intelligent Computing Systems. Chennai, 2024: 1-6.

[24] Wang C Y, Yeh I H, Liao H Y M. YOLOv9: Learning what you want to learn using programmable gradient information[C]. Computer Vision — ECCV 2024. Milan, 2024: 1-21.

[25] Wang A, Chen H, Liu L H, et al. YOLOv10: Real-time end-to-end object detection[C]. Advances in Neural Information Processing Systems 37. Vancouver, 2024: 107984-108011.

[26] Khanam R, Hussain M. YOLOv11: An overview of the key architectural enhancements[PP/OL]. 2024, arXiv: 2410.17725.

[27] Tian Y J, Ye Q X, Doermann D. YOLOv12: Attention-centric real-time object detectors[C]. Advances in Neural Information Processing Systems 38. San Diego, 2025: 87151-87175.

[28] Lei M Q, Li S Q, Wu Y H, et al. YOLOv13: Real-time object detection with hypergraph-enhanced adaptive visual perception[PP/OL]. 2025, arXiv: 2506.17733.

[29] Sapkota R, Cheppally R H, Sharda A, et al. YOLO26: Key architectural enhancements and performance benchmarking for real-time object detection[PP/OL]. 2025, arXiv: 2509.25164.

[30] DJI. DJI Mini 4 Pro, DJI Air 3 and DJI Mavic 3 Pro specifications[EB/OL]. [2026-08-06]. https://www.dji.com/mini-4-pro/specs; https://www.dji.com/air-3/specs; https://www.dji.com/mavic-3-pro/specs.

作者简介

刘佳祺 (1995−), 男, 硕士生, 主要研究方向为低空安全、目标检测, E-mail: 1046849488@qq.com;

张顺香 (1970−), 男, 教授, 博士, 主要研究方向为智能信息处理、情感计算、复杂网络, E-mail: sxzhang@aust.edu.cn;

张正佳 (2000−), 男, 硕士生, 主要研究方向为计算机视觉、目标检测, E-mail: 1009470536@qq.com;

王杰 (1964−), 男, 教授, 博士, 主要研究方向为智能检测与智能仪表、粉尘防治技术, E-mail: 2024095@aust.edu.cn.